

# MASKED FACE RECOGNITION CON IMAGE SIMILARITY

DATASET RANKING E CELEB MASKED SIMULATED

Elio Milite 0522501321  
Giovanni Rapa 0522501352



## INTRODUZIONE

Introduzione al problema

## RELATED WORKS

Diverse soluzioni trattate in letteratura

## SISTEMA PROPOSTO

Dataset Ranking, la nostra rete neurale e Celeb Masked Simulated

01

02

03

# TABELLA DEI CONTENUTI

04

05

## RISULTATI OTTENUTI

Metriche di valutazione della rete sui dataset selezionate e discussione dei risultati

## CONCLUSIONI

Riassunto del lavoro e sviluppi futuri

# 01

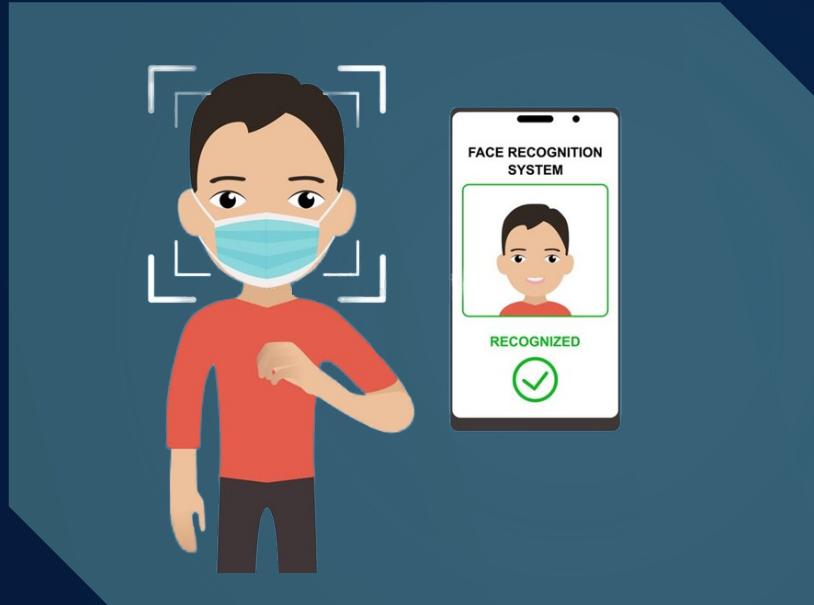
# INTRODUZIONE

---

Cos'è la Masked Face Recognition?



# PERCHÉ MASKED FACE RECOGNITION?



Con il covid-19 e l'introduzione di misure di sicurezza anti-contagio come la mascherina le classiche tecniche di riconoscimento facciale sono diventate obsolete.

La Masked Face Recognition punta a risolvere la problematica col riconoscimento della zona perioculare.



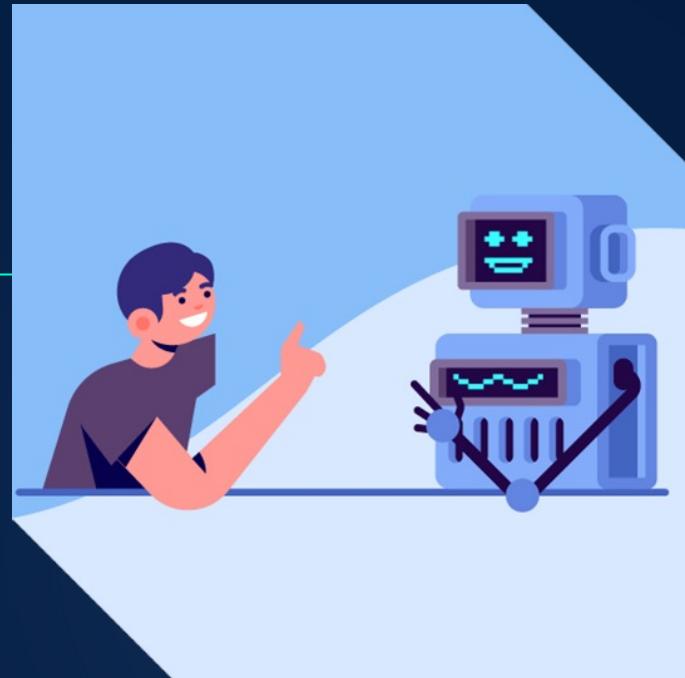


## HUMAN VS MACHINE

---

Lo studio «Masked Face Recognition: Human vs Machine» definisce la difficoltà del riconoscimento facciale, evidenzia la difficoltà nel riconoscere un soggetto con la mascherina e la necessità dell'intervento di un operatore umano.

Sulla falsa riga di questo studio, abbiamo scelto di condurre uno studio approfondito sui dataset e di effettuare dei test per dare un valore all'impatto della mascherina sul volto.



A complex network graph is visible in the background, consisting of numerous small cyan dots connected by thin white lines, forming a dense web of triangles and polygons.

# 02

## RELATED WORKS

---

Come è stato risolto il problema in letteratura?

# MASKED FACE RECOGNITION WITH LATENT PART DETECTION

L'identificazione di volti mascherati rappresenta una sfida per tre motivi principali:

- Non esistono dati di addestramento e di test su larga scala per la Masked Face Recognition.
- Raccogliere e annotare milioni di volti mascherati richiede molto lavoro
- Poiché la maggior parte degli indizi facciali è occlusa dalla maschera, è necessario usare delle rappresentazioni significative riguardanti principalmente la zona perioculare.



Per questo motivo, in questo studio viene introdotto un metodo innovativo per l'incremento dei dati di addestramento per generare automaticamente immagini di volti mascherati.

# TOWARDS NIR-VIS MASKED FACE RECOGNITION



Si fa corrispondere una coppia di immagini facciali acquisite in due modalità diverse come l'infrarosso.

Con una rete neurale semi-siamese, si confrontano i volti effettuando una scansione tridimensionale del volto in analisi.

Inoltre, viene sottolineata la difficoltà data dalla mancanza di informazioni facciali occluse dalla mascherina.

Grazie alla rappresentazione tridimensionale, la soluzione finale fornisce una rappresentazione del volto invariante rispetto al dominio d'applicazione.

# 03

## SISTEMA PROPOSTO

Dataset Ranking, la rete neurale Siamese,  
Celeb Masked Simulated e M2FRED

# DATASET RANKING

# RWMFD: Real World Masked Face Dataset

In questo lavoro vengono proposti 3 tipi di dataset:

- Masked Face Detection Dataset (MFDD)
- Real-world Masked Face Recognition Dataset (RMFRD)
- Simulated Masked Face Recognition Dataset (SMFRD)

Non è stato possibile utilizzare il dataset RMFRD sulla nostra rete perché a file con lo stesso nome corrispondono immagini masked e unmasked di soggetti differenti.

Data la grandezza del dataset non è stato possibile rinominare e trovare le immagini per i giusti confronti.



# MASKED FACE RECOGNITION FOR SECURE AUTHENTICATION

Studio che introduce lo script Mask The Face.

Mask the Face permette di applicare 5 tipi di mascherine a dei volti dati in input.

Il dataset utilizzato per la simulazione con Mask The Face è il Labelled Faces in the Wild (LFW), generalmente utilizzato per misurare le performance di un sistema di riconoscimento facciale.

Esso contiene 5,749 identità per un totale di 13,233 immagini.

Viene inoltre presentato un secondo dataset, chiamato MFR2, contenente foto di celebrità e politici. Anche questo dataset è inutilizzabile per la classificazione perché le identità sono mischiate.





# CELEB MASKED SIMULATED

Data l'impossibilità di classificare correttamente i precedenti dataset, è stato creato un dataset mediante l'uso di Mask The Face.

Celeb Masked Simulated è generato a partire dal dataset MS-CELEB-1M, contenente un milione di immagini di celebrità, dal quale sono state estratte 10.000 identità.

Effettuata l'estrazione, si è applicato Mask The Face, generando 4 immagini con 4 tipi di mascherine diverse per identità.



# CELEB MASKED SIMULATED

Data l'impossibilità di un'attenta classificazione a seguito di un mancato ordinamento, come evidenziato nell'analisi dei dataset precedenti, il nostro dataset si presenta diviso in 5 classi per identità:

- Unmasked: soggetto senza mascherina
- Mask type 1: soggetto avente una mascherina chirurgica bianca.
- Mask type 2: soggetto avente una mascherina di tessuto nera.
- Mask type 3: soggetto avente una mascherina FFP2 blu.
- Mask type 4: soggetto avente una mascherina di tessuto rossa.



# IL DATASET M2FRED

Il dataset M2FRED ci è stato utile per l'estrazione di soggetti con e senza mascherina reale. È stato utilizzato per testare la capacità predittiva del modello su mascherine reali. Dal database contenente video di 43 soggetti, sono stati estratti 200 frame per ogni identità con e senza mascherina reale.



## Tabella di confronto Datasets

Dataset	Anno	Risoluzione	#Soggetti	#Tipi	Tipo	#Immagini
MFDD	2020	256x256	12.000	1	R	24.771
RMFRD	2020	256x256	525	1	R	95.000
SMFRD	2020	256x256	10.000	1	S	500.000
MFR2	2020	160x160	53	1	R	269
LFW-SM	2020	160x160	5.749	4	S	64.973
LFW-DM	2021	403x403	400	2	S	4.117
LMD	2021	Variabile	700	2-3	S	1.510



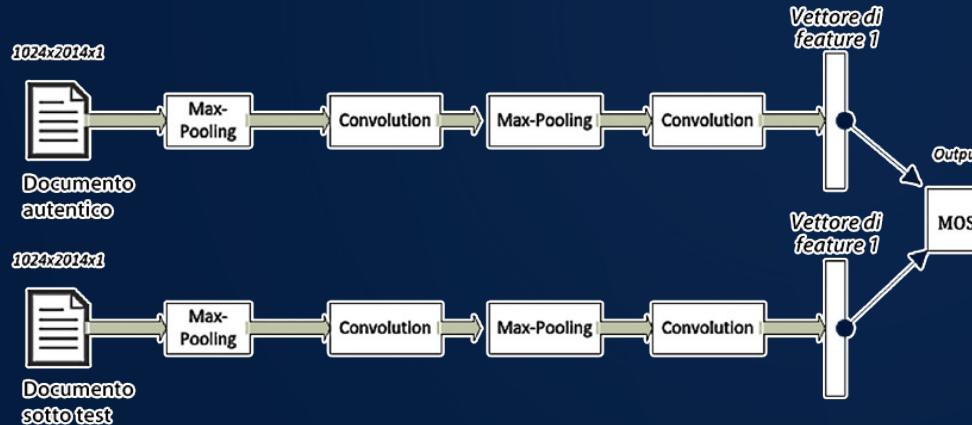
# RETE NEURALE

# LA RETE SIAMESE

Una rete Siamese è una rete neurale formata da due sotto-reti identiche che condividono tra di loro gli stessi pesi.

Permette di massimizzare il contrasto (distanza) tra due input diversi dati alle due reti.

Per il seguente progetto è stata utilizzata per tecniche di image similarity cioè per dare un punteggio alla diversità di due immagini, data dalla mascherina.





# COM'È STATA COSTRUITA?

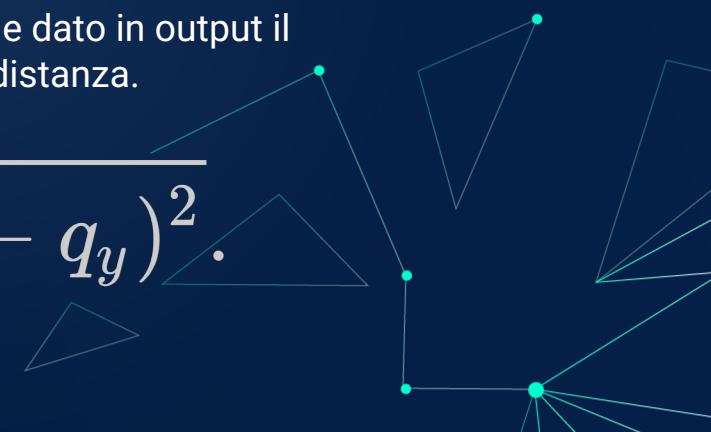
La rete è stata costruita mediante l'uso di TensorFlow e Keras.  
Per valutare effettivamente la differenza tra due immagini è stato  
utilizzato, come input layer, la distanza euclidea.

Essendo una rete di tipo siamese, si utilizzano due input layer  
basati sulla distanza euclidea.

Un layer Lambda unisce i risultati dei due layer utilizzando la  
distanza euclidea e invia il risultato alla rete finale.

Avviene così la valutazione finale e quindi viene dato in output il  
punteggio di somiglianza dato dalla distanza.

$$\sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}.$$





# FUNZIONE DI LOSS: CONSTRANCTIVE

La funzione di perdita descrive l'efficienza del modello rispetto al risultato atteso.  
Esistono vari tipi di funzioni perdita o di costo.

Nella nostra rete neurale è stata implementata la Constranctive loss.

Questa funzione di perdita tratta l'output della rete come un esempio positivo,  
calcola la sua distanza da un esempio della stessa classe e la contrappone alla  
distanza dagli esempi negativi.

$$L(W, Y, \vec{X}_1, \vec{X}_2) =$$

$$(1 - Y)\frac{1}{2}(D_W)^2 + (Y)\frac{1}{2}\{max(0, m - D_W)\}^2$$



# 04

## RISULTATI OTTENUTI

---

Metriche di valutazione della rete  
sui dataset selezionati e  
discussione dei risultati





# METRICHE UTILIZZATE

Sono state usate le seguenti metriche per monitorare le performance del modello:

- Accuracy: è un indicatore sintetico che riassume la capacità del modello di rispondere correttamente.
- Precision: è il rapporto tra il numero delle previsioni corrette di un dato evento sul totale delle volte che il modello lo prevede.
- Recall: misura la sensibilità del modello ed è il rapporto tra le previsioni corrette sul totale dei casi in cui si verifica un evento.
- F1-score: misura l'accuratezza di un test.

Inoltre è stata considerata come metrica la funzione di loss MSE per evidenziarne il grafico ottenuto dato dalla stretta correlazione matematica con la Contrastive loss.



# TEST SU CELEB MASKED SIMULATED



# TEST SU CELEB MASKED SIMULATED



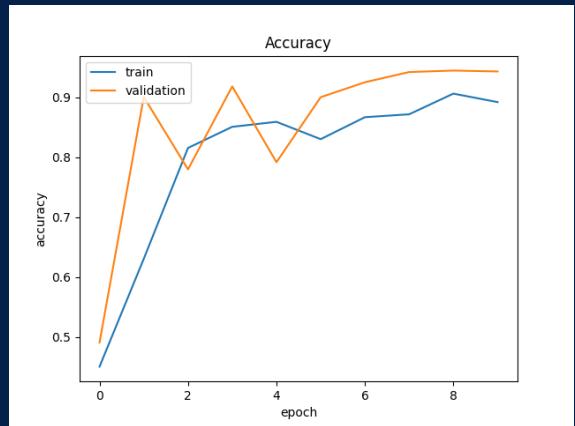
Il dataset è stato diviso in train, validation e test set.

Il modello è stato addestrato su 5 classi: Unmasked, Mask type 1, Mask type 2, Mask type 3, Mask type 4.

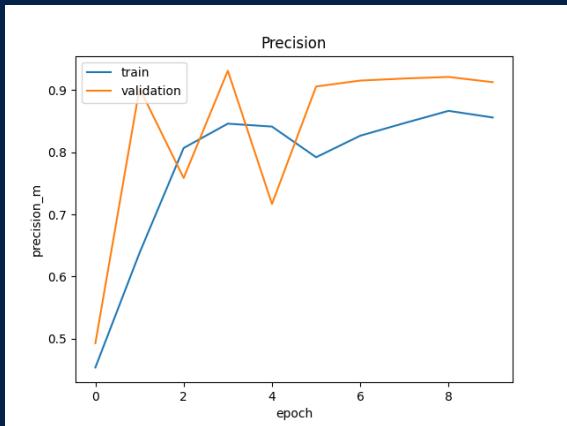
Gli iperparametri utilizzati sono stati settati come segue:

- Epochs = 10
- Batch size = 20
- Margin = 1

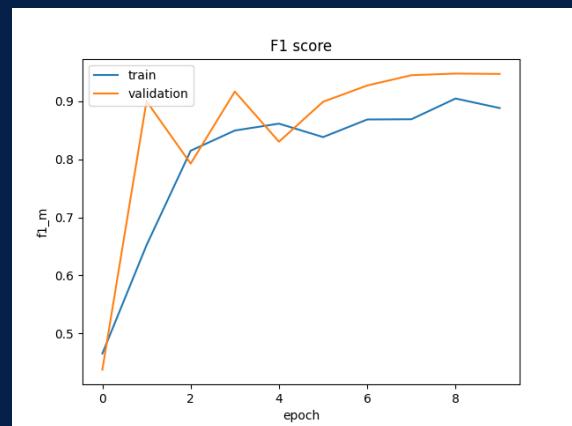
Training: 0.8918. Validation: 0.9430



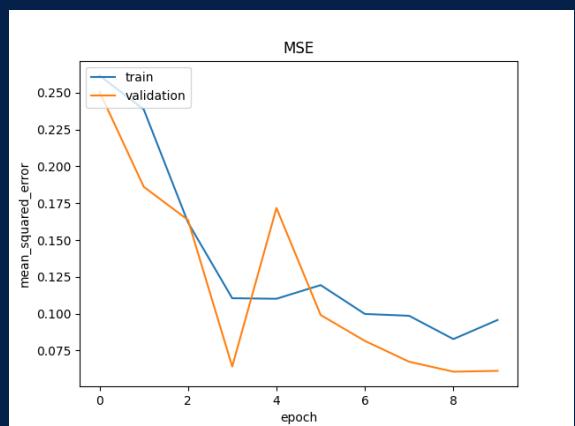
Training: 0.8558. Validation: 0.9126



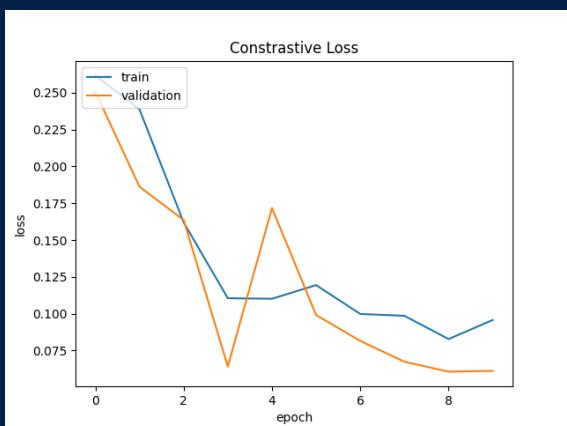
Training: 0.8881. Validation: 0.9468



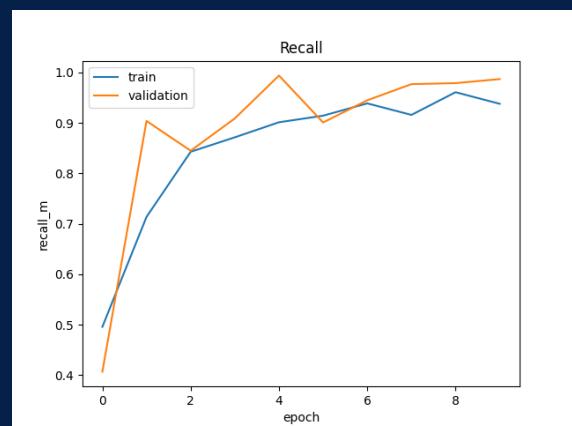
Training: 0.0957. Validation: 0.0612



Training: 0.0957. Validation: 0.0612



Training: 0.9380. Validation: 0.9870

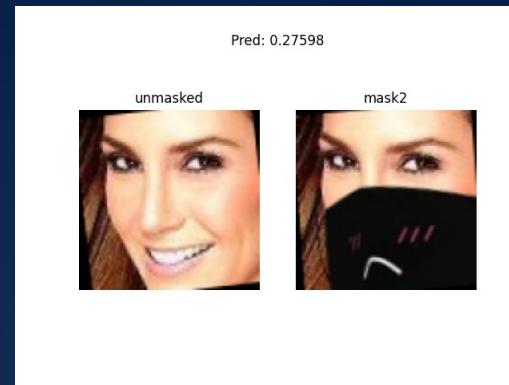


# ALCUNI RISULTATI OTTENUTI

	<b>Mask 1</b>	<b>Mask 2</b>	<b>Mask 3</b>	<b>Mask 4</b>
<b>Identità 1</b>	0.00062	0.27598	0.00037	0.90437
<b>Identità 5</b>	0.00062	0.27598	0.00037	0.90437
<b>Identità 7</b>	0.00062	0.27598	0.00037	0.90437



# Risultati ottenuti



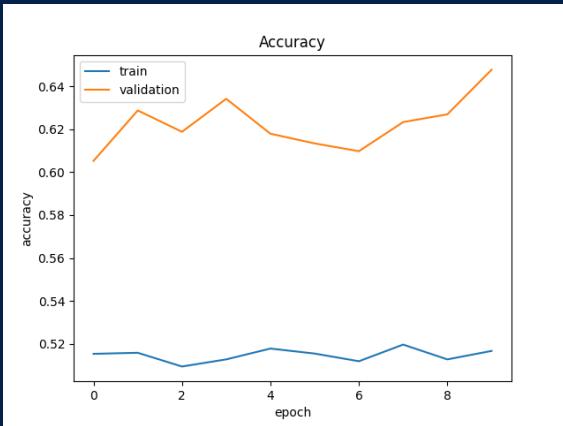


# **SECONDO TEST SU CELEB MASKED SIMULATED**

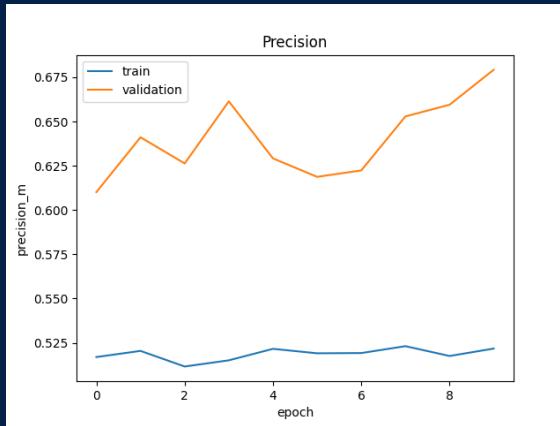
Con il secondo test, il modello è stato addestrato sulle singole identità.  
Ad ogni identità è associata una classe.



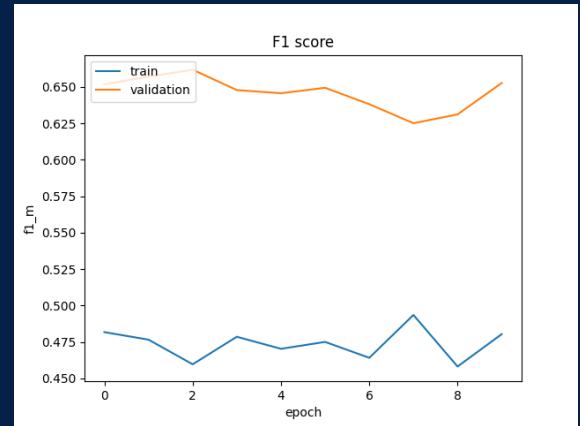
Training: 0.5167. Validation: 0.6477



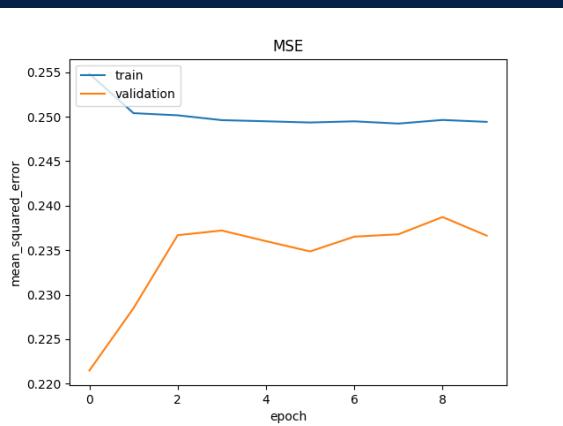
Training: 0.5217. Validation: 0.6792



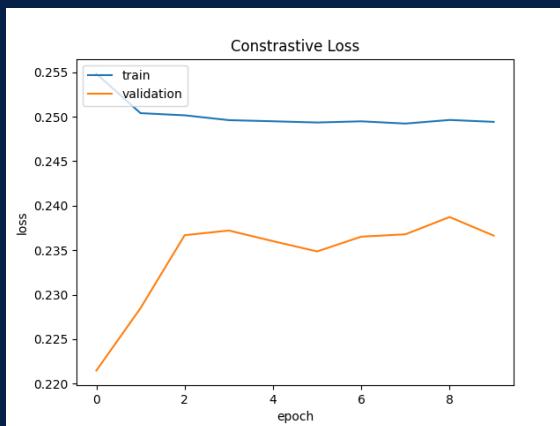
Training: 0.4804. Validation: 0.6527



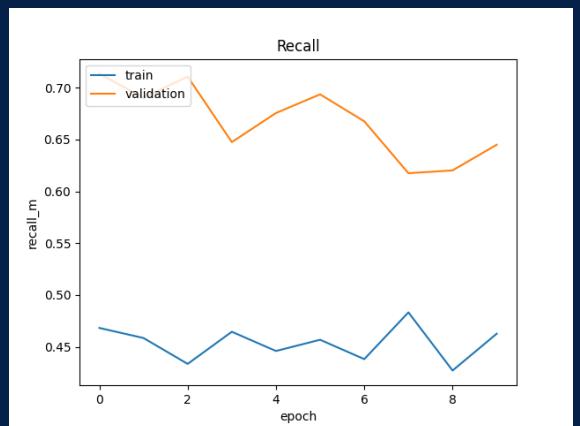
Training: 0.2366. Validation: 0.2494



Training: 0.2494. Validation: 0.2366



Training: 0.4628. Validation: 0.6449

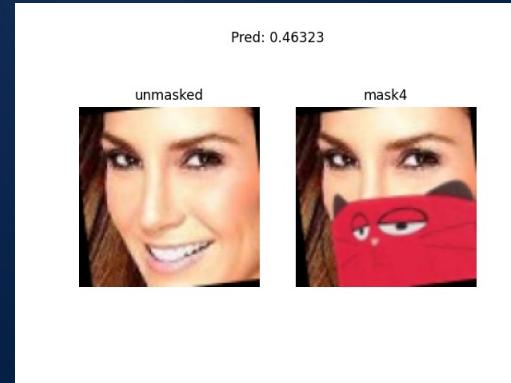
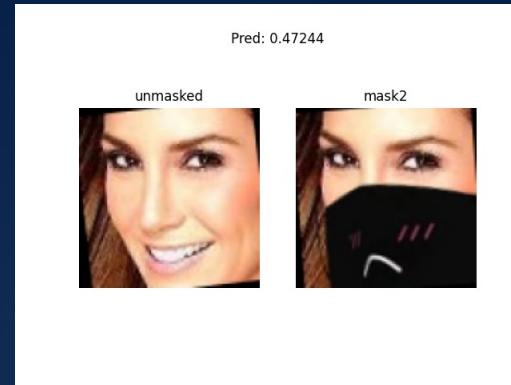
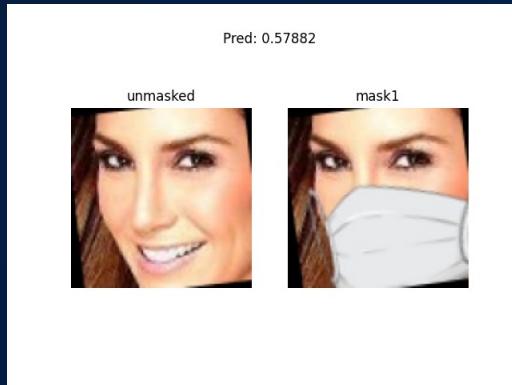


# ALCUNI RISULTATI OTTENUTI

	<b>Mask 1</b>	<b>Mask 2</b>	<b>Mask 3</b>	<b>Mask 4</b>
<b>Identità 1</b>	0.57882	0.47244	0.55560	0.46323
<b>Identità 5</b>	0.50853	0.46733	0.47291	0.53250
<b>Identità 7</b>	0.53250	0.46133	0.46733	0.52003



# Risultati ottenuti



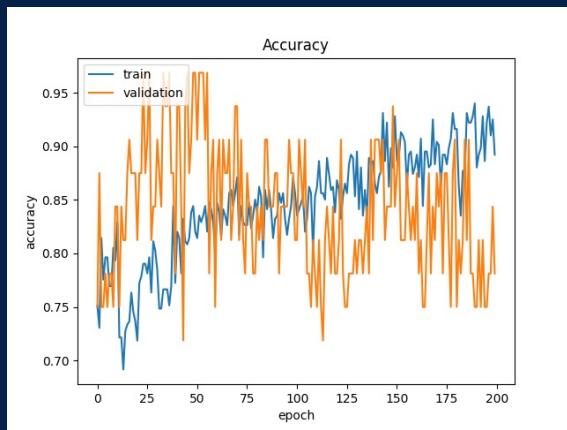


# TEST SU M2FRED

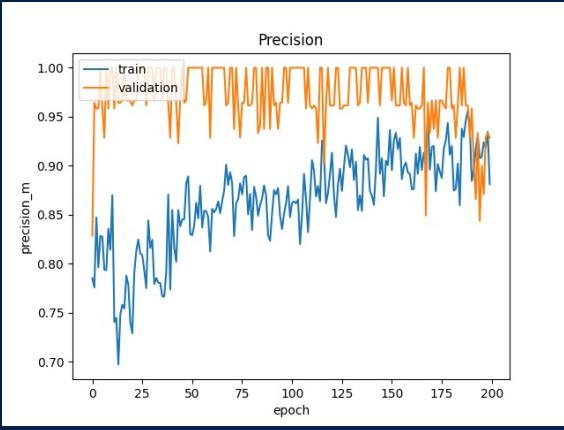
con 200 epochhe



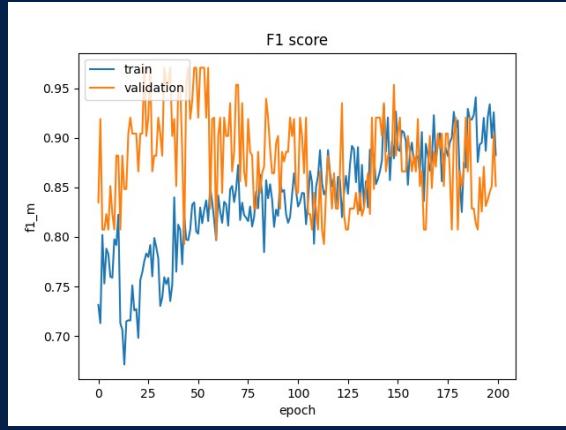
Training: 0.8918. Validation: 0.7812



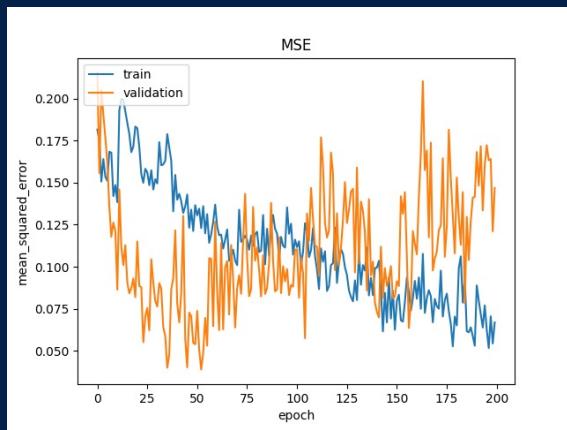
Training: 0.8812. Validation: 0.9286



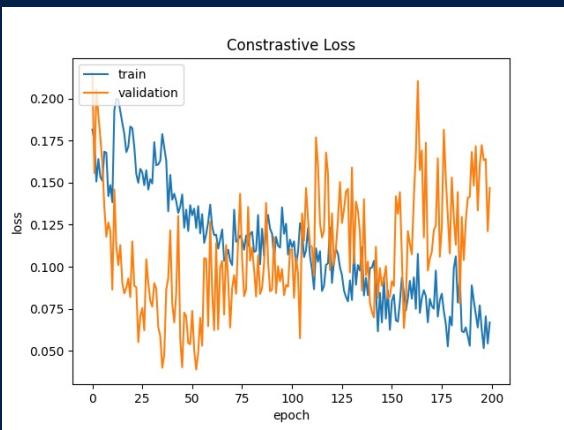
Training: 0.8826. Validation: 0.8513



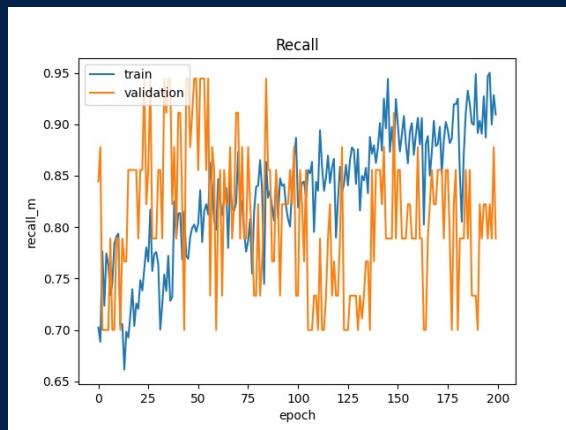
Training: 0.0668. Validation: 0.1469



Training: 0.0668. Validation: 0.1469



Training: 0.9094. Validation: 0.7889



# ALCUNI RISULTATI OTTENUTI

	Session 1	Session 2	Session 3	Session 4
Identità 1	0.12	0.52	0.83	0.83
Identità 3	0.59	0.83	0.22	0.32
Identità 4	0.28	0.36	0.80	0.92



# RISULTATO MIGLIORE VS RISULTATO PEGGIORE

Pred: 0.12080

001



001



VS

Pred: 0.92353

004



004



# 05

# CONCLUSIONI

Riassunto del lavoro e sviluppi futuri



## LAVORI FUTURI

Per continuare sulla falsa riga di questo lavoro, sarà necessario ampliare il dataset da noi proposto così da poter ri-addestrare una generica rete neurale per migliorarne le capacità predittive.

Proponiamo anche di cambiare gli iper-parametri per cercare una configurazione migliore sulla propria rete neurale.





# **GRAZIE A TUTTI PER L'ATTENZIONE**

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), and infographics & images by [Freepik](#).

Please keep this slide for attribution.